网络攻击检测的门控记忆网络方法 *

王家宝，徐伟光，周振吉，李　阳，苗　壮

(陆军工程大学，南京 210007)

摘　要：针对互联网大规模网络攻击检测难题，结合词向量特征表示与循环神经网络，提出了一种门控记忆网络检测方法。该方法首先将网络请求数据转换为低维实值向量序列表示，然后利用门控循环神经网络的长时记忆能力提取请求数据的特征，最后采用逻辑斯特回归分类器实现了对网络攻击的自动检测。在 CSIC2010 公开数据集上，达到了 98.5% 的 10 折交叉验证 F1 分数。与传统方法相比，较大幅度地提高了网络攻击检测的准确率和召回率。所提方法可自动检测网络攻击，具有良好的检测效果。

关键词：网络攻击检测；低维实值向量表示；门控循环神经网络

中图分类号：TP393.08　　doi: 10.3969/j.issn.1001-3695.2018.01.0169

Gated memory network approach for Web attack detection

Wang Jiabao, Xu Weiguang, Zhou Zhenji, Li Yang, Miao Zhuang

(*Army Engineering University of PLA*, *Nanjing* 210007, *China*)

Abstract: To solve the problem of large-scale network attack detection, this paper proposed a gated memory network method, based on word vector feature representation and recurrent neural network. Firstly, the proposed method transformed the network request data into low-dimension real-value vector sequence representation. And then, it extracted the features of request data by using the memory ability of gated recurrent neural network. Finally, it adopted the logistic regression classifier to achieve automatic detection of network attack. On the CSIC2010 public data set, this method achieves 98.5% 10-fold cross-validation F1-score. Comparing with traditional methods, it can effectively improve the precision and recall rates for detecting network attack. The proposed method can detect network attacks automatically and has good detection results.

Key words: network attack detection; low-dimension real-value vector representation; gated recurrent neural networks

0　引言

随着互联网规模的不断发展，网络攻击已成为各国安全部门和企业面临的重大问题。由于网络攻击手段多样、类型各异，且极易变种，导致对网络攻击的检测也面临着巨大困难。针对互联网大规模的网络通信行为，如果能够自动地判断通信行为的恶意性，则可以有效地避免网络攻击行为产生的破坏及其可能的次生灾害。目前，针对网络攻击行为的检测主要分为基于模式匹配的检测和基于模式分类的检测[1]。前者是当前大多数网络安全软件的主要手段，即通过模式匹配或统计分析判断系统或网络日志中的不正常行为，如登录不期望的位置、访问未授权文件、网络流量异常、程序行为异常等[2]。该类检测方法通常是在攻击行为发生或事后进行的检测，无法在事前和事中进行预先检测判断。后者通常先可对通信内容提取特征[3]，再利用分类器直接进行检测。该类方法可直接针对网络通信流数据进行，对发现的异常通信数据直接处理或丢弃，避免攻击行为对主机造成影响。

基于模式分类的检测方法是本文研究的重点。目前该方法主要是借鉴文本分类技术，通过提取内容的描述特征，将网络攻击检测问题转换为模式分类问题进行处理。其中特征表示方法主要包括词袋（Bag-of-Words）表示[4]、TF-IDF 表示[5]、n-gram 表示[6]；分类检测方法主要包括贝叶斯[7]、决策树[8]、支持向量机[9,10]、多层感知机[11,12]、K-近邻[13]等经典机器学习方法。

近年来，随着深度学习的发展，特征表示与分类识别研究内容均得到了新发展。在文本特征表示方面，谷歌研究组提出的 word2vec 模型能够将传统高维的 one-hot 词向量表示转换为低维实值向量表示，改变了文本词的表达能力，使得词与词之间可以度量距离（相似度）[14,15]。该特征表示在情感分类方面

收稿日期：2018-01-30；**修回日期：**2018-04-16　　**基金项目：**国家重点研发计划基金资助项目（2017YFB0802900）

作者简介：王家宝（1985-），男，安徽肥西人，讲师，博士，主要研究方向为机器学习、计算机视觉（jiabao_1108@163.com）；徐伟光（1984-），男，安徽宿州人，讲师，博士，主要研究方向为机器学习、网络安全防护；周振吉（1985-），男，江苏连云港人，讲师，博士，主要研究方向为网络攻击与防护；李阳（1984-），男，河北廊坊人，讲师，博士研究生，主要研究方向为计算机视觉、机器学习；苗壮（1976-），男，副教授，博士，主要研究方向为智能信息处理.

得到了很好的应用[16]，但主要限于与文本词内容的相关研究。对于网络通信而言，基于网络协议的通信请求是由字符串组成的流序列，可以被转换为文本词序列以低维实值向量序列进行表示。同时，对于产生的序列数据，循环神经网络（recurrent neural network, RNN）和门控循环神经网络（gated RNN）具有良好的建模能力，并在文本分类[17]、入侵检测[18]、机器翻译[19]等领域得到了广泛应用。

受低维实值词向量模型[14,15]和门控循环网络模型[19,20]的启发，本文提出了一种网络攻击检测的门控记忆网络方法。该方法中，基于 HTTP 网络通信的请求内容被表示为低维实值向量序列，增强了数据的特征表示能力；同时，具有长时记忆能力的门控循环神经网络模型被用于建模序列中长间隔的词间关系，有效提升了分类检测效果。与传统方法相比，在 CSIC2010 公开数据集上，所提门控记忆网络方法较大幅度地提高了网络攻击检测的准确率和召回率。

1　网络攻击检测的门控记忆网络方法

1.1　门控记忆网络方法的流程框架

图 1 是所提出网络攻击检测的门控记忆网络方法的流程框架。其中，攻击用户通过 Web 协议发出攻击请求，该请求数据是网络攻击检测的原始数据。

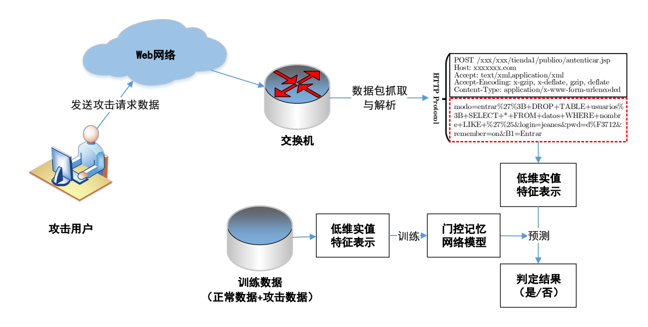


图 1　网络攻击检测的门控记忆网络方法流程框架

该流程框架中，首先从网络交换机抓取并解析数据包；然后解析后数据，经过低维实值特征表示转变为门控记忆网络模型的输入数据；最后通过预先训练的门控记忆网络模型对未知类别的输入数据进行预测，判定是攻击请求或正常请求。门控记忆网络模型的参数是在训练数据上预先学习得到的。训练数据包含正常数据和攻击数据两类样本。经过低维实值特征表示输入模型进行训练。该流程框架的核心内容是低维实值特征表示和门控记忆网络模型。

1.2　低维实值特征表示

网络攻击的实施依赖于网络通信，且多以异常的通信请求出现。通常，通信请求可以被看做是一个命令字符串，而异常的网络通信请求数据中通常包含特殊的命令字符串，如 systemInfo、alert、SELECT 等，因此可以通过对请求数据中包含的字符串进行分类来检测网络攻击。

图 2 是基于 HTTP 网络协议攻击数据的低维实值特征表示过程。该过程主要包括词切分、量化和降维三个步骤。

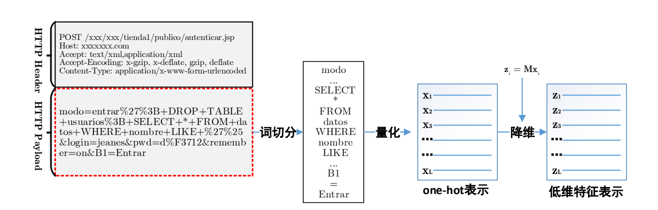


图 2　HTTP 网络协议攻击数据的低维实值特征表示过程

1）词切分　对于通信请求字符串而言，其内容是根据网络协议的规则构造的，故其格式是具有一定规范的，也是能够被转换成由一系列“单词”或符号组成的文本。例如，针对网络服务器的访问很大一部分是基于 HTTP 协议的，而请求数据的载荷部分中所提交的参数字符串，可以视为由符号“&”分隔的若干段，每个段内存在键和值两部分，以“=”连接。因此，可以将请求字符串分隔成若干“单词”或符号的序列，即使“单词”或符号是没有字面意义。在此基础上，借助现有的文本表示技术实现对请求数据的量化表示。

2）量化　在词切分之后，传统文本分类的特征表示方法通常将每个词表示为一个 one-hot 向量。one-hot 向量是一个二进制向量，在词汇表中第 i 个单词的第 i 个元素被设置为 1，其他元素都被设置为 0，以此唯一地表示一个词。对于一个由 L 个词组成的序列而言，可以表示为一个长度为 L 的 one-hot 序列 $(x_1, x_2, ..., x_L)$。但是，one-hot 向量是一个稀疏高维二值向量，该特征表示在计算是耗时较大，且难以度量两个词之间的距离（相似度）关系，如两个词是语义相近的同义词。

3）降维　为了克服 one-hot 向量表示的不足，本文将一个 one-hot 向量 x_i 投影为一个低维空间中的实值向量 $z_i \in \mathbb{R}^d$（d 为空间维度）。该投影过程可以通过对向量 x_i 左乘一个投影矩阵 $M \in \mathbb{R}^{d \times |V|}$ 实现（式（1）），其中 $|V|$ 是无重复词典中词的个数。

$$z_i = Mx_i \tag{1}$$

矩阵 M 可以通过随机赋值得到或通过包含一个隐层的网络学习得到。实验表明，通过学习得到的矩阵 M 具有更好的低维表示。矩阵 M 学习的网络输入一个 one-hot 词向量，输出与其相邻的下一个 one-hot 向量词，以此来学习两个共现词之间的关系。网络训练完成后，其隐藏输出即为低维实值向量 z_i。降维表示将一个高维稀疏的 one-hot 序列 $(x_1, x_2, ..., x_L)$ 转换为一个低维实值向量序列 $(z_1, z_2, ..., z_L)$。

1.3　门控记忆网络模型

门控记忆网络模型由一个门控循环神经网络和一个逻辑斯特回归分类器组成。其中，门控循环神经网络对序列数据进行建模，抽取长时记忆特征表示；逻辑斯特回归分类器完成对请求数据的二类分类预测判定。

1.3.1 门控循环神经网络

近年来，循环神经网络在语音识别领域取得了优越的成果，其结构主要由一个循环过程构成：

$$h_t = \varphi(W_h z_t + R_h h_{t-1} + b_h) \tag{2}$$

其中：t 时刻的激活 h_t 由 t 时刻的输入 z_t 和 $t-1$ 时刻的激活 h_{t-1}

决定，如图 3 左侧所示。循环神经网络通过循环迭代计算 $h_{t-1} \to h_t$ 使信息可以得到长时记忆，对于序列数据具有优秀的建模能力。但是经典的循环神经网络在训练模型时面临着梯度弥散问题[21]。为了克服该问题，长短期记忆（long short-term memory, LSTM）单元结构[18,22]被引入循环神经网络。该结构由输入门、遗忘门、输出门等门控单元组成，以维持长时序列数据的信息记忆。由于网络由一系列门控单元组成，也被称为门控循环神经网络。

受门控循环神经网络[19,20]的启发，网络攻击检测的门控记忆网络单元结构如图 3 右侧所示。其中，z_t 为低维实值特征表示的输入向量；h_{t-1} 为网络 $t-1$ 时刻的隐状态；h_t 为网络 t 时刻输出的隐状态；h_t 不仅作为 $t+1$ 时刻的输入，还作为计算最终输出结果的输入。虚线框中，箭头表示数据流向；xx 表示点乘操作；++ 表示向量加操作；1-1– 表示对输入 x 变量 1-x 操作；σ 表示 sigmoid 激活操作；tanh 表示 tanh 激活操作。低维实值向量序列 $(z_1, z_2, ..., z_L)$ 进入门控记忆网络后会按序列展开进行计算，当最后一个向量 z_L 进入网络经计算后，对应输出 h_L 即为门控记忆网络的输出。

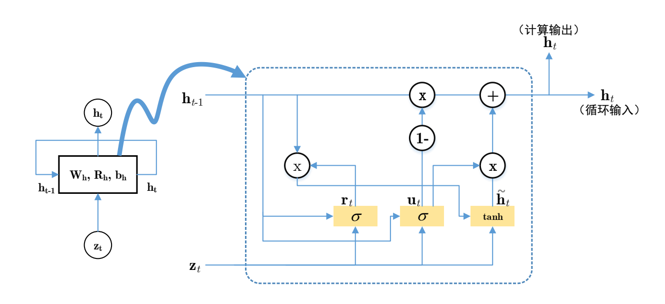


图 3　循环神经网络(左)与门控记忆网络单元结构(右)

图 3 中门控记忆网络单元结构主要由重置门和更新门组成，形式化如下：

$$r_t = \sigma(W_r z_t + R_r h_{t-1} + b_r) \tag{3}$$

$$u_t = \sigma(W_u z_t + R_u h_{t-1} + b_u) \tag{4}$$

其中：W_r, W_u 与 R_r, R_u 分别为输入 z_t 和隐状态 h_{t-1} 的权重；b_r, b_u 为对应的偏置；$\sigma(\cdot)$ 为 sigmoid 函数。最终输出：

$$h_t = (1-u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t , \tag{5}$$

其中：$\tilde{h}_t = \varphi(W_h z_t + r_t \odot (R_h h_{t-1}) + b_h)$；$\varphi(\cdot)$ 为 tanh 函数；$\odot$ 为元素点乘；W_h、R_h 和 b_h 分别为输入和隐状态的权重，以及对应偏置。$\tilde{h}_t$ 的计算结果受重置门控制，若 $r_t = 0$，则 $r_t \odot (R_h h_{t-1}) = 0$，即 $t-1$ 时刻的信息输入 $R_h h_{t-1}$ 不起作用，此时隐含信息 $\tilde{h}_t$ 仅由 t 时刻输入信息 z_t 控制，即 $\tilde{h}_t$ 遗忘了其历史信息 h_{t-1}，被 t 时刻输入信息 z_t 重置。最终输出 h_t 是由 u_t 控制加权平均 $t-1$ 时刻隐含信息 h_{t-1} 与 t 时刻隐含信息 $\tilde{h}_t$ 得到。

与 LSTM 相比，门控记忆网络单元结构将输入门和遗忘门整合为更新门，以平衡 $t-1$ 时刻的激活 h_{t-1} 和 t 时刻的更新激活 $\tilde{h}_t$，同时重置门通过元素点乘决定是否遗忘 $t-1$ 时刻的激活 $R_h h_{t-1}$。序列数据的最终输出为具有长时记忆特性的隐含状态 h_L，该向量作为分类器的输入。

1.3.2 逻辑斯特回归分类器

对于网络攻击检测二类分类问题，输入为网络协议请求数据，经低维实值特征表示转换为长度为 L 的向量序列 $(z_1, z_2, ..., z_L)$；再经门控循环神经网络提取长时记忆特征 h_L；最终通过一个二类分类器进行分类输出。

本文采用逻辑斯特回归分类器。训练时，给定一组二值标签样本 $\{(h_L^{(i)}, y^{(i)}) : i = 1, ..., N\}$，其中：$h_L^{(i)}$ 为第 i 个样本的特征向量；$y^{(i)} \in \{1, 0\}$ 为其对应的真实类别标签（1 表示攻击样本，0 表示正常样本）；N 为训练样本个数。根据最大似然估计原理，计算所有样本的似然如下：

$$L = \prod_i \left(\mathcal{F}_\theta(h_L^{(i)})\right)^{y^{(i)}} \left(1 - \mathcal{F}_\theta(h_L^{(i)})\right)^{1-y^{(i)}} \tag{6}$$

其中：$\mathcal{F}_\theta(v) = 1/(1+\exp(-\theta^T v))$ 逻辑斯特函数，其值可表示对输入样本 v 判定为攻击样本的概率；θ 为待学习参数向量。为了方便优化，将最大化 L 转变为最小化负的对数似然：

$$\begin{aligned} J &= -\log L \\ &= -\log \prod_i (\mathcal{F}_\theta(h_L^{(i)}))^{y^{(i)}} (1 - \mathcal{F}_\theta(h_L^{(i)}))^{1-y^{(i)}} \\ &= -\sum_i \left(\log(\mathcal{F}_\theta(h_L^{(i)}))^{y^{(i)}} + \log(1 - \mathcal{F}_\theta(h_L^{(i)}))^{1-y^{(i)}}\right) \\ &= -\sum_i \left(y^{(i)} \log(\mathcal{F}_\theta(h_L^{(i)})) + (1-y^{(i)}) \log(1 - \mathcal{F}_\theta(h_L^{(i)}))\right) \end{aligned} \tag{7}$$

对于一个训练样本而言，式（7）中求和的两项只有一项不为零（取决于标签 $y^{(i)}$ 是 1 或 0）。当 $y^{(i)} = 1$，最小化优化目标 J 意味着需要使 $\mathcal{F}_\theta(h_L^{(i)})$ 变大；当 $y^{(i)} = 0$，则需要使 $1 - \mathcal{F}_\theta(h_L^{(i)})$ 变大。

实际过程中，逻辑斯特回归分类器可与门控记忆网络连接起来共同训练，门控记忆网络输出隐状态 $h_L^{(i)}$ 通过一个一个全连接层来计算输出 $q = \theta^T h_L^{(i)}$，再连接一个 sigmoid 层实现 $o = 1/(1+\exp(-q))$ 的计算，最后损失层计算式（7）结果。优化时，采用随机梯度下降算法进行参数更新，从后向前先计算偏导数 $\partial J / \partial o$，再根据反向传播算法，依次计算各层的偏导数以实现网络参数的学习。

测试时，将损失函数替换为一个逻辑斯特函数 $\mathcal{F}_\theta(v)$。当需要预测判定一个新的样本是属于“1”还是属于“0”时，则可以通过比较 $\mathcal{F}_\theta(h_L^{(i)})$ 与 $1 - \mathcal{F}_\theta(h_L^{(i)})$ 的大小来进行判定。若 $\mathcal{F}_\theta(h_L^{(i)}) > 1 - \mathcal{F}_\theta(h_L^{(i)})$，则判为“1”，否则判为“0”。

1.4　门控记忆网络方法的实现细节

对于 CSIC2010 数据集，先从中提取的通信请求数据。对于 GET 请求直接提取 URI 信息，对于 PUT 和 POST 请求提取 URI 和负载数据并将两者拼接起来，作为低维实值特征表示的输入，经过低维实值特征表示转换为一个词向量序列。具体实现时，词向量序列由一个词索引序列和一个词索引到词向量的映射矩阵组成。词索引序列中每个索引对应一个预训练好的词向量，根据词索引可从词索引到词向量的映射矩阵中找到词向量，该表示可大大节省空间。词索引到词向量的查找由一个网络嵌入层（embedding layer）实现。值得注意的是，词向量序列会根据数据集的整体特征被统一截断或补齐为长度 56 的序列，

目的是可以批量输入数据进行训练。

门控记忆网络方法采用了深度学习的架构体系，其具体框架如表 1 所示。

表 1　门控记忆网络方法框架

网络层	描述
输入层	输入词索引序列，不做处理直接输出
嵌入层	输入词索引序列和词索引到词向量的映射矩阵，将词索引转化为词向量输出
门控记忆层	输入词向量序列，门控记忆单元计算的最后一个隐状态作为输出
Dropout 层	输入隐状态，以一定概率丢弃权值后输出
全连接层	通过全连接映射到两个值输出
Softmax 层	输出两个数值，归一化为两个概率值输出

表中，输入层和嵌入层完成低维实值特征表示；Dropout 层是深度学习中用于抑制过拟合的主要手段，使用该策略可以获得更好的测试精度；全连接层和 Softmax 层实现逻辑斯特回归，对应计算式（6）中的 $\mathcal{F}_\theta(h_L^{(i)})=1/(1+\exp(-\theta^T h_L^{(i)}))$，全连接完成线性映射计算，Softmax 对映射后的两个值归一化得到 $\mathcal{F}_\theta(h_L^{(i)})$ 和 $1-\mathcal{F}_\theta(h_L^{(i)})$。训练时，网络计算损失并进行梯度回传和参数更新；测试时，网络计算 $\mathcal{F}_\theta(h_L^{(i)})$ 和 $1-\mathcal{F}_\theta(h_L^{(i)})$ 大小，判定类别标签。

基于 Windows 7 操作系统（3.5 GHz 主频，8 GB 内存）和 tensorflow 平台，实现门控记忆网络方法，所有训练和测试均采用 CPU 完成。网络输入为 d 维的向量序列，单层门控记忆单元的隐状态维度为 128，其后网络接一个 Dropout 层，Dropout 率为 0.9，然后将 128 维输出全连接映射到 2 维结果输出。训练网络时，将批处理大小设置为 128，学习率设置为 0.001。训练代数为 10，使用 Adam 优化器训练。

2　实验结果与分析

实验以 CSIC2010 数据集作为测试数据集。该数据集包含上万条自动生成的 HTTP 协议请求，主要用于测试网络攻击防护系统，它是由西班牙研究委员会（CSIC）信息安全研究所制作的。该数据集针对的是一个电子商务 Web 应用程序，已发布的数据被分为训练（只有正常）和测试（异常和正常）集。在实验中，采用的是测试集中 36 000 多个正常请求和 25 000 多个异常请求。该数据集中的攻击请求包含多种网络攻击，如 SQL 注入、缓冲区溢出、信息收集、文件披露，CRLF 注入、跨站脚本和参数窜改等。其中，针对隐藏（或不可用）资源的请求也被视为异常[1,14]。目前，由于个人隐私保护等原因，公开可用的 Web 攻击检测问题的数据集非常少，诸如 DARPA KDD99 攻击检测数据集中的攻击很多都已过时，且不包括许多新型的攻击类型。

对于原始的请求数据，实验主要提取 GET、POST、PUT 请求数据来进行检测。请求数据提取后对数据进行分词，即字符串分割。分词依据 HTTP 请求特点进行，主要涉及 URL 中字符解码，参数项、键值对、特殊符号的分割，请求数据的分词可以为后续低维实值特征表示提供基础。

2.1　方法对比

为了验证本文方法的有效性，分别与传统方法和深度学习方法进行对比分析。传统方法包括朴素贝叶斯（naïve Bayes, NB）、线性支持向量机（linear support vector machine, LSVM）、神经网络（neural network, NN）、K-近邻（K-near neighbor, KNN）、决策树（decision tree, DT）。深度学习方法包括 LSTM。

在特征表示方面，传统方法统一采用 TF-IDF 特征向量。深层学习方法采用 1.2 节中描述的数据特征低维表示。在分类检测方面，传统方法的参数设置如下：NB 恒定的参数为 0.01。LSVM 采用线性分类器。NN 采用两个隐层，隐单元个数分别为 50 和 10。KNN 设置邻居参数为 3。DT 采用熵最小，且叶子中的最小样本数被设置为 3。

所有方法采用 10 折交叉验证技术进行结果评估。数据集随机分为 10 份，其中一份用于测试，其余用于训练。10 次运行的结果取平均得到整体性能。测试结果以准确率、召回率和 F1 分数作为评测指标。实验结果见表 2。

表 2　实验训练和测试数据

方法	准确率	召回率	F1 分数
本文方法	**0.984**	**0.985**	**0.985**
LSTM	0.977	0.979	0.978
DT	0.925	0.917	0.920
KNN	0.907	0.911	0.908
NN	0.895	0.882	0.887
LSVM	0.887	0.867	0.875
NB	0.767	0.776	0.765

由表 2 可以发现，门控记忆网络方法明显优于其他方法，准确率和召回率分别达到了 98.4%和 98.5%，较 LSTM 超出 0.7%的准确率和 0.6%的召回率，且远超传统方法。该结果表明本文方法具有很好的攻击检测效果。

2.2　隐变量参数分析

门控记忆网络中，隐状态维度是影响网络能力的关键参数因子。为了进一步分析该参数变化对性能的影响。本节中固定 Dropout 参数为 0.9，低维实值词向量维度为 40。测试不同维度条件下，模型检测方法的 F1 分数综合性能。为了进一步说明本文方法的有效性，同时对比不同参数条件下 LSTM 方法的 F1 分数，实验结果如图 4 所示。此外，不同参数条件下，模型训练耗时如表 3 所示。

表 3　模型训练耗时/min

方法\隐变量维度	16	32	64	128	256	512
本文方法	17	24	40	58	132	416
LSTM	17	24	36	67	175	539

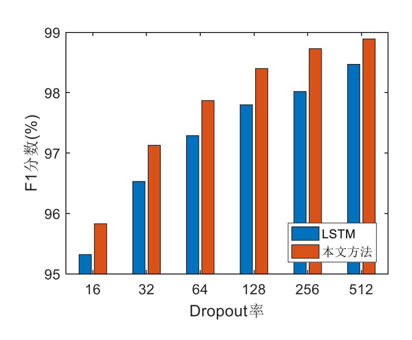


图 4　不同隐节点维度下的 F1 分数

由图 4 和表 3 可以得到如下结论：

a）模型的性能随着参数维度的增加而不断增加，但增加的幅度越来越小；

b）对比 LSTM 模型，可知本文方法在相同参数条件下较 LSTM 模型结果更好；

c）模型计算耗时会随着维数的增加而增加，故实际应用需要在效果与性能之间进行折中。

2.3　Dropout 策略分析

考虑到 CSIC2010 数据集的规模中等，为了抑制学习的过拟合问题，在模型训练中，加入了 Dropout 策略。该策略中 Dropout 率决定了模型的效果。故研究固定隐变量维度为 128，低维实值词向量维度为 40 的条件下，不同 Dropout 率的影响。图 5 给出了不同方法的 Dropout 率的结果。

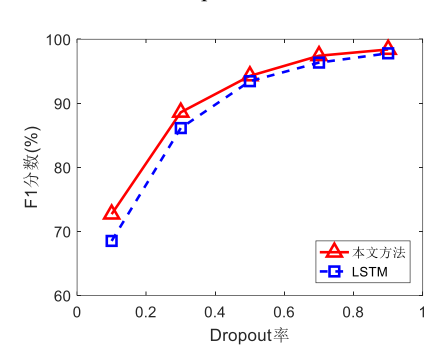


图 5　不同 Dropout 率下的 F1 分数

由图 5 可以得到：

a）模型泛化性能随着 Dropout 率增加而不断增长，但增长幅度不断减小。

b）Dropout 策略（较大的 Dropout 率）具有很好的抑制过拟合的能力，并取得了较好的性能。Dropout 率为 0.9 时，达到了 98.40%的 F1 分数。

2.4　低维实值词向量分析

低维实值词向量具有很好的特征表示能力，可以度量不同词间的相似性关系。具有相似关系或相同属性的词汇在向量空间中距离更近，易聚集在一起。为了说明低维实值词向量的表示能力，本文选取部分低维实值词向量并通过 t-SNE 方法[23]嵌入二维平面显示词间的关系，如图 6 所示。

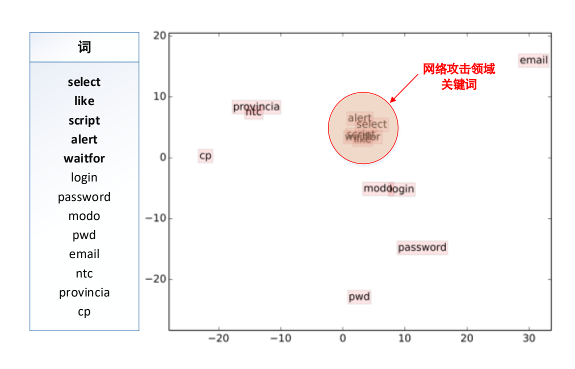


图 6　低维实值词向量嵌入的分布

图 6 给出了 13 个低维实值词向量的嵌入显示。其中 select、like、script、alert、waitfor 是 SQL 入侵攻击的关键词，这些词在嵌入空间中均聚集在一起，而正常词汇如 login、password 等则呈现随机分布的特性，这种相似词聚集的特性更容易帮助分类器学习网络攻击行为的模式。

此外，不同的低维实值词向量维度也会对模型的精度产生影响。为了研究低维实值词向量维度对算法性能的影响，分别选取 10、20、40、80、160、320，测量模型对攻击检测的效果和耗时。图 7 给出了对应的测试结果。

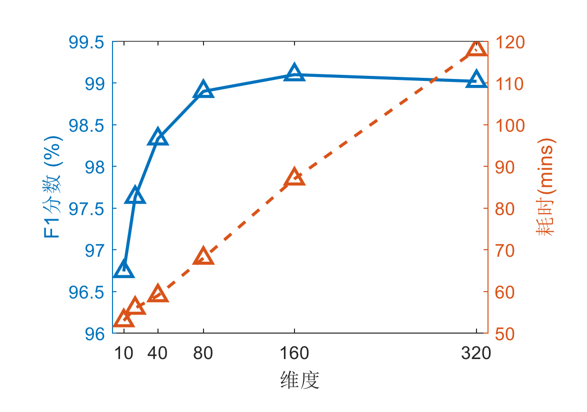


图 7　低维实值词向量不同维度下的 F1 分数和耗时

由图 7 可知：

a）本文方法随着低维实值词向量维度的增加，检测效果也不断增加，在维度为 160 时，达到了 99.10%的 F1 分数；

b）同时，本文方法的耗时会随低维实值词向量维度的增加不断增长，且基本保持线性增长速率；

c）在实际应用中，可根据应用的差异选择一定的低维实值词向量维度以实现速度和效果的折中。

3　结束语

针对网络攻击检测问题，本文提出了一个门控记忆网络方法，在 CSIC2010 数据集上达到了 98.40%的 F1 分数，超越了传统方法和 LSTM 等检测方法。该方法实现简单、效果显著。同时，算法在检测时速度可达到实时应用需要。但是，由于数据特征低维表示的学习通常需要一定规模的数据支撑，所以本文需要大量网络请求数据来进行学习。

此外，本文方法仅采用单层门控循环神经网络结构，下一

步拟采用多层网络结构进一步提高检测精度。考虑到多层网络的时间耗费，可根据实际情况构造两层或三层网络结构，以折中检测精度和时间耗费。同时，为了进一步验证本文的效果，后续将在实际网络环境中进行测试，以检验方法的实际可用性。

参考文献：

[1] Elbachirelmoussaid N, Toumanari A. Web application attacks detection: a survey and classification [J]. International Journal of Computer Applications, 2014, 103 (12): 1-6.

[2] 徐周波，张永超，古天龙，等. 面向入侵检测系统的模式匹配算法研究 [J]. 计算机科学, 2017, 44 (9): 125-130. (Xu Zhoubo, Zhang Yongchao, Gu Tianlong, *et al.* Research on pattern matching algorithm in intrusion detection system [J]. Computer Science, 2017, 44 (9): 125-130.)

[3] 戴远飞，陈星，陈宏，等. 基于特征选择的网络入侵检测方法 [J]. 计算机应用研究, 2017, 34 (8): 2429-2433. (Dai Yuanfei, Chen Xing, Chen Hong, *et al.* Feature selection based approach to network intrusion detection [J]. Application Research of Computers, 2017, 34 (8): 2429-2433.)

[4] 马林进，万良，马绍菊，等. 基于词袋模型的分布式拒绝服务攻击检测 [J]. 计算机应用, 2017, 37 (6): 1644-1649. (Ma Linjin, Wan Liang, Ma Shaoju, *et al.* Distributed denial of service attack recognition based on bag of words model [J]. Journal of Computer Application, 2017, 37 (6): 1644-1649.)

[5] Mao Chinghao, Lee H M, Liu Ensi, *et al.* Web mimicry attacks detection using HTTP token causal correlation [J]. International Journal of Innovative Computing Information & Control, 2011, 7 (7): 4347-4362.

[6] Oza A, Ross K, Low R M, *et al.* HTTP attack detection using n-gram analysis [J]. Computers & Security, 2014, 45 (3): 242-254.

[7] Singh K J, De T. An approach of DDOS attack detection using classifiers [M]// Emerging Research in Computing, Information, Communication and Applications. New Delhi: Springer, 2015: 429-437.

[8] García V H, Monroy R, Quintana M. Web attack detection using ID3 [M]// Professional Practice in Artificial Intelligence. Boston, MA: Springer, 2006: 323-332.

[9] Rawat R, Kumar Shrivastav S. SQL injection attack detection using SVM [J]. International Journal of Computer Applications, 2012, 42 (13): 1-4.

[10] 吴少华，程书宝，胡勇. 基于SVM的Web攻击检测技术 [J]. 计算机科学, 2015, 42 (6A): 362-364. (Wu Shaohua, Cheng Shubao, Hu Yong. Web attack detection method based on support vector machines [J]. Computer Science, 2015, 42 (6A): 362-364.)

[11] Silva L D S, Santos A C F D, Silva J D S D, *et al.* A neural network application for attack detection in computer networks [C]// Proc of IEEE International Joint Conference on Neural Networks. Budapest, . Hungary: IEEE Press, 2004: 1569-1574.

[12] Li Jin, Liu Yong, Gu Lin. DDoS attack detection based on neural network [C]// Proc of IEEE International Symposium on Aware Computing. [S. l.] : IEEE Press, 2010: 196-199.

[13] 肖甫，马俊青，黄洵松，等. SDN环境下基于KNN的DDoS攻击检测方法 [J]. 南京邮电大学学报：自然科学版, 2015, 35 (1): 84-88. (Xiao Fu, Ma Junqing, Huang Xunsong, *et al.* DDoS attack detection based on KNN in software defined networks [J]. Journal of Nanjing University of Posts and Telecommunications: Natural Science Edition, 2015, 35 (1): 84-88.)

[14] Mikolov T, Chen K, Corrado G, *et al.* Efficient estimation of word representations in vector space [J/OL]. CoRR, 2013, abs/1301. 3781. https://arxiv. org/abs/1301. 3781.

[15] Mikolov T, Sutskever I, Chen K, *et al.* Distributed representations of words and phrases and their compositionality [C]// Proc of 27th Annual Conference on Neural Information Processing Systems. Lake Tahoe, Nevada: Curran Associates, Inc, 2013: 3111-3119.

[16] Nowak J, Taspinar A, Scherer R. LSTM recurrent neural networks for short text and sentiment classification [C]// Proc of International Conference on Artificial Intelligence and Soft Computing. Cham: Springer, 2017: 553-562.

[17] Liu Pengfei, Qiu Xipeng, Huang Xuanjing. Recurrent neural network for text classification with multi-task learning [C]// Proc of International Joint Conference on Artificial Intelligence. New York, NY: IJCAI//AAAI Press, 2016: 2873-2879.

[18] Kim J, Kim J, Thu H L T, *et al.* Long short term memory recurrent neural network classifier for intrusion detection [C]// Proc of IEEE International Conference on Platform Technology and Service. Jeju, SouthKorea: IEEE Press, 2016: 1-5.

[19] Cho K, Merrienboer B V, Gulcehre C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation [J/OL]. CoRR, 2014, abs/1406. 1078. https://arxiv. org/abs/1406. 1078.

[20] Chung J, Gulcehre C, Cho K H, *et al.* Empirical evaluation of gated recurrent neural networks on sequence modeling [J/OL]. CoRR, 2014, abs/1412. 3555. https://arxiv. org/abs/1412. 3555

[21] Pascanu R, Mikolov T, Bengio Y. On the difficulty of training recurrent neural networks [C]// Proc of International Conference on Machine Learning. Altanta, Georgia: PMLR, 2013: 1310-1318.

[22] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.

[23] Maaten L V D, Hinton G. Visualizing data using t-SNE [J]. Journal of Machine Learning Research, 2008, 9 (2605): 2579-2605.